

1 Binding letter of intent as advance notification of a full renewal proposal

x	Binding letter of intent (required as advance notification for renewal proposals in 2025)
---	---

2 Formal details

- Name of the consortium

FAIRmat – FAIR Data Infrastructure for Condensed-Matter Physics, and the Chemical Physics of Solids

- Acronym of the consortium

FAIRmat

- Applicant institution

Humboldt-Universität zu Berlin
Unter den Linden 6, 10117 Berlin
Prof. Dr. Julia von Blumenthal

- Spokesperson

Claudia Draxl
claudia.draxl@physik.hu-berlin.de
Humboldt-Universität zu Berlin

- Co-applicant institutions

Leibniz-Institut für Kristallzüchtung (IKZ)
Max-Born-Straße 2, 12489 Berlin
Prof. Dr. Thomas Schröder

Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

Freyeslebenstraße 1, 91058 Erlangen
Prof. Dr. Joachim Hornegger

Ruhr-Universität Bochum (RUB)

Universitätsstraße 150, 44801 Bochum
Prof. Dr. Martin Paul

Helmholtz-Zentrum Berlin (HZB)

Hahn-Meitner-Platz 1, 14109 Berlin

Prof. Dr. Bernd Rech

Max-Planck-Gesellschaft e.V.**Max Planck Computing and Data Facility (MPCDF)**

Gießenbachstraße 2, 85748 Garching

Prof. Dr. Erwin Laure

FAIR Data Infrastructure for Physics, Chemistry, Materials Science, and Astronomy e.V. (FAIR-DI)

Zum Großen Windkanal 2, 12489 Berlin

Prof. Dr. Claudia Draxl

▪ Co-spokespersons

Martin Albrecht

martin.albrecht@ikz-berlin.de

Leibniz-Institut für Kristallzüchtung (IKZ)

Christoph T. Kochchristoph.koch@hu-berlin.de

Humboldt-Universität zu Berlin (HU)

Heiko B. Weber

heiko.weber@fau.de

Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

Silvana Botti

silvana.botti@rub.de

Ruhr-Universität Bochum (RUB)

Erwin Laureerwin.laure@mpcdf.mpg.de

Max Planck Computing and Data Facility (MPCDF)

Kevin Jablonka

kevin.jablonka@helmholtz-berlin.de

Helmholtz-Zentrum Berlin (HZB)

Michael Krieger

michael.krieger@fau.de

FAIR Data Infrastructure for Physics, Chemistry, Materials Science, and Astronomy e.V. (FAIR-DI) / Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU)

▪ Participant institutions

Deutsche Physikalische Gesellschaft (DPG)

Hauptstraße 5, 53604 Bad Honnef

TU Dresden - Zentrum für Informationsdienste und Hochleistungsrechnen (ZIH)

Zellescher Weg 12-14, 01069 Dresden

Technische Universität München (TUM)

Arcisstraße 21, 80333 München

Bayerische Akademie der Wissenschaften

Leibniz-Rechenzentrum (LRZ)

Boltzmannstraße 1, 85748 Garching bei München

Helmholtz-Zentrum Hereon

Max-Planck-Straße 1, 21502 Geesthacht

▪ Participant individual

Martin Aeschlimann	Rheinland-Pfälzische Technische Universität Kaiserslautern-Landau
Hilke Bahmann	Bergische Universität Wuppertal
Tristan Bereau	Universität Heidelberg
Oliver Bierwagen	Paul-Drude-Institut für Festkörperelektronik
Maik Bierwirth	Humboldt-Universität zu Berlin (HUB)
Sandor Brockhauser	Max-Planck-Institut für Kognitions- und Neurowissenschaften
Hans-Joachim Bungartz	Technische Universität München (TUM)
Keith Butler	University College London, UK
Fabio Caruso	Christian-Albrechts-Universität zu Kiel
Cecilia Clementi	Freie Universität Berlin (FUB)

Deutsche Forschungsgemeinschaft

Kennedyallee 40 · 53175 Bonn, Germany · Postal address: 53170 Bonn, Germany
Tel.: + 49 228 885-1 · Fax: + 49 228 885-2777 · postmaster@dfg.de · www.dfg.de

Andrea Crovetto	Danmarks Tekniske Universitet, Denmark
Felix Dietrich	Technische Universität München
Malte Dreyer	Humboldt-Universität zu Berlin
Natascha Dropka	Leibniz-Institut für Kristallzüchtung
Ralph Ernstorfer	Technische Universität Berlin
Claudia Felser	Max-Planck-Institut für Chemische Physik fester Stoffe
Katharina Franke	Freie Universität Berlin
Pascal Friederich	Karlsruher Institut für Technologie (KIT)
Angela Fösel	Friedrich-Alexander-Universität Erlangen-Nürnberg
Janine George	Bundesanstalt für Materialforschung und -prüfung (BAM) / Friedrich-Schiller-Universität Jena
Luca Ghiringhelli	Friedrich-Alexander-Universität Erlangen-Nürnberg
Roger Gläser	Universität Leipzig
Vincenzo Grillo	CNR Istituto Nanoscienze, Modena, Italy
Axel Groß	Universität Ulm
Marius Grundmann	Universität Leipzig
Thomas Hammerschmidt	Ruhr-Universität Bochum
Tamás Haraszi	DWI – Leibniz-Institut für Interaktive Materialien
Stefan Hecht	Humboldt-Universität zu Berlin
Thomas Heine	Technische Universität Dresden
Hendrik Heinz	University of Colorado Boulder, USA
Kerstin Helbig	Humboldt-Universität zu Berlin
Alex Henderson	The University of Manchester, UK
Walid Hetaba	Max-Planck-Institut für Chemische Energiekonversion
Oliver Hilt	Ferdinand-Braun-Institut (FBH)
Stefan Kesselheim	Forschungszentrum Jülich (FZJ)
Thomas Kosch	Humboldt-Universität zu Berlin
Angelina Kraft	Leibniz-Informationszentrum Technik und Naturwissenschaften und Universitätsbibliothek (TIB)
Peter Kraus	Technische Universität Berlin (TUB)
Christopher Künneth	Universität Bayreuth
Thomas Lunkenbein	Universität Bayreuth
Miguel A. L. Marques	Ruhr-Universität Bochum
Knut Müller-Caspary	Ludwig-Maximilians-Universität München (LMU)
Wolfgang E. Nagel	Technische Universität Dresden
Nina Owschikow	Technische Universität Berlin
Uli Paetzold	Karlsruher Institut für Technologie
Rossitza Pentcheva	Universität Duisburg-Essen
Laurenz Rettig	Fritz-Haber-Institut der Max-Planck-Gesellschaft
Carsten Richter	Leibniz-Institut für Kristallzüchtung
Gian-Marco Rignanese	University of Louvain, Belgium
Patrick Rinke	Technische Universität München
Raphael Ritz	Max Planck Computing and Data Facility

Giorgio Rossi	Università degli Studi di Milano, Italy
Mariana Rossi	Max-Planck-Institut für Struktur und Dynamik der Materie
Erich Runge	Technische Universität Ilmenau / Deutsche Physikalische Gesellschaft (DPG)
Barbara Scalise	Leibniz-Institut für Kristallzüchtung
Thomas Schröder	Leibniz-Institut für Kristallzüchtung
Julia Schumann	Leibniz-Institut für Polymerforschung Dresden
Jonathan Scragg	Uppsala University, Sweden
Hélène Seiler	Freie Universität Berlin
Erdmann Spiecker	Friedrich-Alexander-Universität Erlangen-Nürnberg
Helge S. Stein	Technische Universität München
Simon Stier	Fraunhofer-Institut für Silicatiforschung
Karoline Stolze	Leibniz-Institut für Kristallzüchtung
Huayna Terraschke	Christian-Albrechts-Universität zu Kiel
Milica Todorović	University of Turku, Finland
Annette Trunschke	Fritz-Haber-Institut der Max-Planck-Gesellschaft
Eva Unger	Helmholtz-Zentrum Berlin
Thomas Unold	Helmholtz-Zentrum Berlin
Denis Usvyat	Humboldt-Universität zu Berlin
Roser Valenti	Goethe-Universität Frankfurt am Main
Aron Walsh	Imperial College London, UK
Holger von Wenckstern	Universität Leipzig
Marco Werner	Leibniz-Institut für Polymerforschung Dresden
Stefan Wesner	Universität zu Köln
Julia Westermayr	Universität Leipzig
Francois Willaime	French Alternative Energies and Atomic Energy Commission (CEA), France
Angela Wittmann	Johannes Gutenberg-Universität Mainz
Bernd Witzigmann	Friedrich-Alexander-Universität Erlangen-Nürnberg
Sabine Wurmehl	Leibniz-Institut für Festkörper- und Werkstoffforschung Dresden
Ursula Wurstbauer	Universität Münster
Hongbin Zhang	Technische Universität Darmstadt
Dan Zhou	Leibniz-Institut für Kristallzüchtung

3 Objectives, work programme and research environment in the second funding period

- Research area of the proposed consortium (according to the DFG classification system: www.dfg.de/subject-classification)

FAIRmat represents the fields of **condensed-matter physics** (DFG Review Board 3.21) and **the chemical physics of solids** (DFG Review Board 3.12).

- Concise summary of the consortium's main objectives and task areas

In the first funding period, FAIRmat has elaborated a rich portfolio of concepts and technologies, that now allow a consistent workflow for research data management (RDM), bringing already an immediate benefit for the colleagues who adapt to our concepts. So far, we have organized our work with a central technology service – the NOMAD ecosystem – and specifications for pilot fields within the large community of condensed matter physics.

Science in this field has changed in recent years. It is by now more than evident that the upcoming paradigm of data-centric research will become of utmost importance for our community. FAIRmat has anticipated this in the first proposal, and has already proven to be a major source of AI-centric science. However, like in any other scientific domain, breakthroughs in AI-supported research remain limited by the lack of highly structured, interoperable data. This insight reemphasizes the necessity of most advanced RDM concepts and an even broader cultural change in the entire field. In the second funding period, FAIRmat adapts to this paradigm change and sets much more ambitious goals. FAIRmat anticipates that highly structured data and interoperability will be a strategic prerequisite for the 4th paradigm of science. In this mindset, we have adapted the consortium's mission and structure.

We now briefly outline the main technological and conceptual modules that collectively shape the FAIRmat strategy and its success. A core element is the comprehensive data management platform NOMAD (see also Figure 1). Being an infrastructure that puts defined data structure and interoperability as primary design principles, it fosters the definition of rich data schemas, acquisition from a wide range of sources, transformation of raw files into structured and reusable data, data sharing across institutions, and scalable analysis.

Of utmost importance for the collaborative culture of our scientific community is the ability to connect distributed and federated infrastructure through a network of local NOMAD instances, known as NOMAD Oases. NOMAD Oasis is an internationally adopted software used to organize

data for individual small research groups, entire universities, and large-scale collaborations, as well as industrial research units. It includes electronic lab notebook (ELN) functionality and enables users to search, access, and analyze data inside the chosen environment. In parallel, the central NOMAD installation acts as an open-science community-wide data hub and as connection node for (meta)data contributed by local instances. NOMAD CAMELS is an experimental control software that lets users orchestrate their lab with a continuously growing library of instrument drivers thanks to the support of the community. It provides further a simple entrance pathway into the NOMAD ecosystem.

The original approach of developing interoperability concepts through selected pilot methods has led to a profound understanding of the requirements and solutions. In the next funding period, we can leave such separations behind and aim for overarching concepts, without sacrificing method specificity. We maintain **Synthesis**, **Experiment**, and **Computation** as vertical task areas, because they ensure the bottom-up approach by our PIs bringing in their invaluable domain expertise from different research fields. Within these domains, we generalize the so far distinct methods and emphasize on the overarching systematics, allowing us to address much larger fractions of our vast community and their challenges.

Building on the already mature concepts elaborated in the first funding period, we introduce a layered structure of three horizontal areas – **Data Modeling and Interoperability**, **Digital Infrastructure**, and **Enabling Data-centric Science**. This picture emphasizes that (meta)data are the prerequisite for AI being successful in our domain, and that the infrastructure provides a performant linking element as the technical backbone of our ecosystem.

Our dedicated **Outreach** task area organizes the manifold interactions between the FAIRmat team and the broader community. Its focus slightly shifts to stronger harmonization with other consortia (OneNFDI) and collaborations with international stakeholders. The FAIRmat Hub at the Center for the Science of Materials Berlin (CSMB) will remain the primary working environment of the core team. It is managed and administered by a rather lean Head Office (task area **Management**). This centralized approach has proven highly effective in building and maintaining infrastructure at scale and will continue to serve as a cornerstone of FAIRmat's operations in the second funding period.

- Brief description of the proposed use of existing infrastructures, tools and services that are essential in order to fulfil the planned consortium's objectives

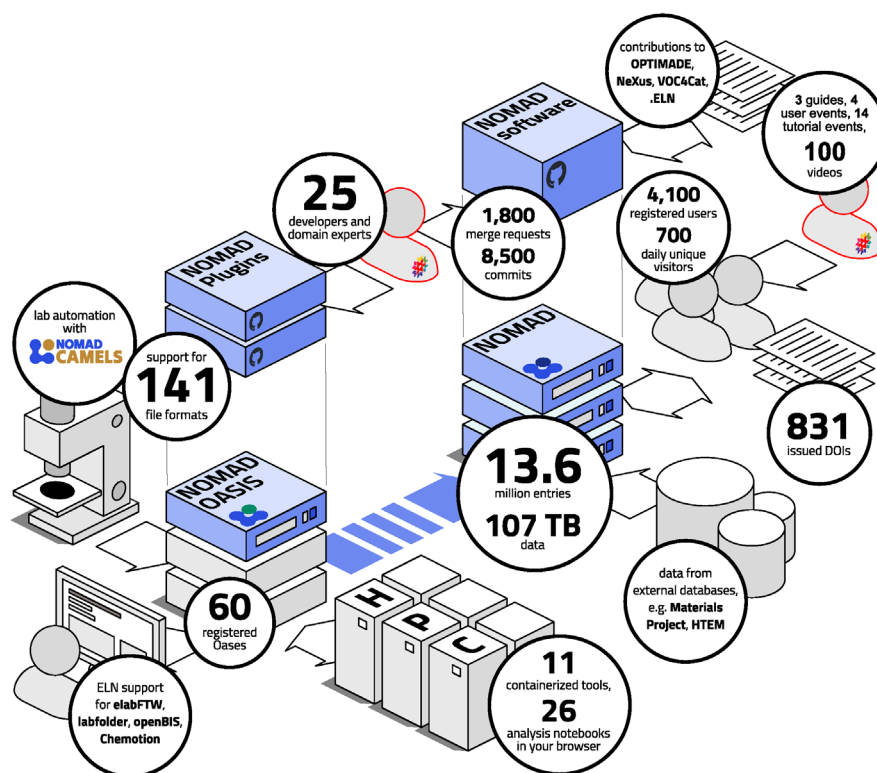


Figure 1: NOMAD, the integrated, extensible, and distributed RDM ecosystem.

FAIRmat will continue developing the NOMAD ecosystem into an even more powerful and intuitive data platform. A diagram illustrating FAIRmat's current services is shown in Figure 1 with the key components and performance indicators of the NOMAD ecosystem as of February 2025.

The central NOMAD services, including various test, staging, and production NOMAD deployments, test Oases, the AI Toolkit and other JupyterHub instances, as well as our AAI service are operated at our co-applicant institution MPCDF, the Max Planck Data and Computing Facility. A large portion of FAIRmat's software development is done on their GitLab instance. However, the majority of data and software services runs on local Oases, which are serviced by FAIRmat members and beneficiaries on their own resources. So far, 80 NOMAD Oases are registered, many of them in environments beyond FAIRmat's core community. Note, however, that we are aware of many more. Currently, the infrastructure is well adapted to the needs, but the increasing demands from several communities may impose bottlenecks.

- Interfaces to other NFDI consortia: brief description of existing agreements for collaboration and/or plans for future collaboration

Collaboration is at the heart of FAIRmat. FAIRmat not only engages within its disciplinary domain, but also aims at connecting different consortia and supporting interdisciplinary integration. As such, FAIRmat has initiated the Physical Sciences in NFDI (PSinNFDI), a collaborative forum that includes DAPHNE4NFDI, MaRDI, NFDI-MatWerk, NFDI4Cat, NFDI4Chem, and PUNCH4NFDI. With most of our neighbors, joint projects are already running and are planned to be intensified:

- NeXus development and linking to related ontologies with DAPHNE4NFDI, bridging small labs and large-scale facilities;
- Common outreach activities with DAPHNE4NFDI;
- Common vocabulary with NFDI4Cat;
- Integration of Chemotion in NOMAD with NFDI4Chem;
- Offering quantum-chemistry methods in NOMAD with NFDI4Chem;
- Pyiron integration into NOMAD with NFDI-MatWerk;
- Shared tools and metadata standards for electron microscopy and atom probe tomography with NFDI-MatWerk;
- Experience exchange and common strategies for teaching with DAPHNE4NFDI, PUNCH4NFDI, and NFDI4Chem;
- Represented in Jupyter4NFDI, PID4NFDI, and IAM4NFDI with Base4NFDI;
- Contributor to AAI incubator projects with Base4NFDI;

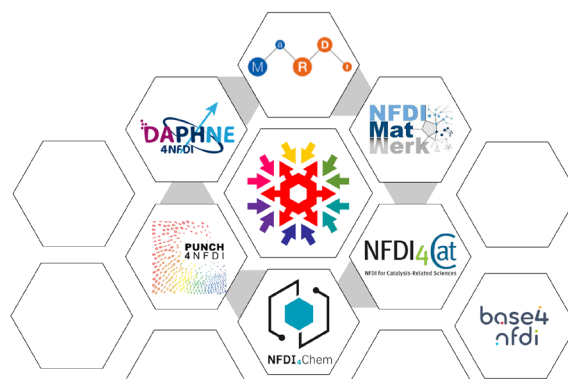


Figure 2: FAIRmat's neighboring consortia with ongoing and planned collaborative activities.

Besides strengthening our integration with these core services, we plan to deepen our involvement in the upcoming terminology service and foster further harmonization between domain-specific and overarching RDM infrastructures.

4 International and national networking

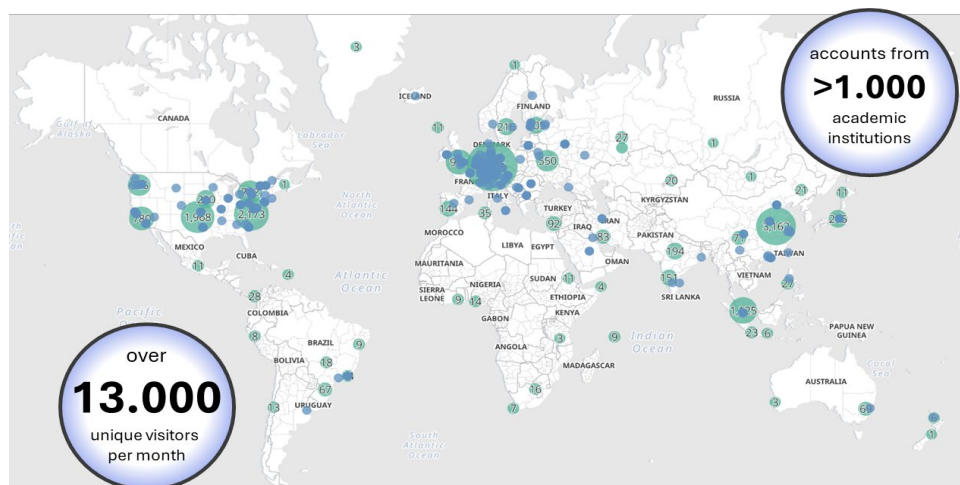


Figure 3: Global distribution of the unique visitors of NOMAD's central service.

At any stage, FAIRmat's infrastructure was an international data platform. This is reflected in the reach of NOMAD's central service, which now has over 13,000 unique monthly visitors. There is particularly strong engagement in Germany as a result of NFDI support. Although registration is not needed for download, NOMAD has over 5,000 registered users. We will work towards its full integration into the EOSC federation as part of the NFDI node. In the upcoming funding period, FAIRmat will further strengthen its international footprint, with new PIs joining from institutions in Denmark, France, Finland, Belgium, Italy, Sweden, and the UK. Also in several European projects, we will continue offering our services and support to the broader European research community such as CHARISMA, AddMorePower, EUSpecLab, MaMMoS, and the COST Actions DAEMON, and QUAST and upcoming ones. Another ongoing project is SolMates (Horizon Europe). We will keep our close collaborations with UK's Physical Science Data Infrastructure program and with U.S. national labs, including NREL, Oak Ridge National Lab, and the Molecular Foundry at Lawrence Berkley National Lab. We will also deepen our outreach in Asia, building on existing collaborations with MatDX.org in Korea and RDM activities and organizations in Japan.

FAIRmat members have long played an active role in shaping global RDM standards and will continue to do so, through leadership positions in the Research Data Alliance (RDA). FAIRmat will continue extending the NeXus data standards and partnering with international photon and neutron facilities, including ALBA, DIAMOND, and ESRF. Within the OPTIMADE consortium, FAIRmat will maintain its role in defining interoperable APIs for materials databases, NOMAD being the largest OPTIMADE contributor.